

5288549601

09/785,435



INVESTOR IN PEOPLE

The Patent Office  
Concept House  
Cardiff Road  
Newport  
South Wales  
NP10 8QQ

RECEIVED  
JUN 06 2001  
Technology Center 2600

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

I also certify that the attached copy of the request for grant of a Patent (Form 1/77) bears a correction, effected by this office, following a request by the applicant and agreed to by the Comptroller-General.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed

Dated

29 December 2000

CERTIFIED COPY OF  
PRIORITY DOCUMENT

**THIS PAGE BLANK (USPTO)**



The  
Patent  
Office

1/77  
29DEC99 E501729-1 D02829  
P01/7700 0.00-779318.7  
The Patent Office

# Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

Cardiff Road  
Newport  
Gwent NP10 8QQ

1. Your reference	J00041975GB		
2. Patent application number (The Patent Office will fill in this part)	9930518.7 23 DEC 1999		
3. Full name, address and postcode of the or of each applicant (underline all surnames)	Mitsubishi Electric Information Technology Centre Europe BV 20 Frederick Sanger Road The Surrey Research Park Guildford, Surrey GU2 5YD United Kingdom  Patents ADP number (if you know it)  If the applicant is a corporate body, give the country/state of its incorporation United Kingdom Netherlands 4028/7/00 41775/7/00		
4. Title of the invention	METHOD AND APPARATUS FOR TRANSMITTING A VIDEO IMAGE		
5. Name of your agent (if you have one)	R G C JENKINS & CO  "Address for service" in the United Kingdom to which all correspondence should be sent (including the postcode) 26 Caxton Street London SW1H 0RJ  Patents ADP number (if you know it) 950001✓		
6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications (and if you know it) the or each application number	Country	Priority application number (if you know it)	Date of filing (day / month / year)
7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application	Number of earlier application		Date of filing (day / month / year)
8. Is a statement of inventorship and of right to grant of a patent required in support of this request (Answer 'Yes' if: a) any applicant named in part 3 is not an inventor, or b) there is an inventor who is not named as an applicant, or c) any named applicant is a corporate body. See note (d))	YES		

9. Enter the number of sheets for any of the following items you are filing with this form.  
Do not count copies of the same document

Continuation sheets of this form

Description	13
Claim(s)	6
Abstract	1
Drawing(s)	4 <del>tt</del> <i>ft</i>

10. If you are also filing any of the following, state how many against each item.

Priority documents

Translation of priority documents

Statement of inventorship and right to grant of a patent (Patents Form 7/77)

Request for preliminary examination and search (Patents Form 9/77) 1

Request for substantive examination (Patents Form 10/77)

Any other documents (please specify)

11. I/We request the grant of a patent on the basis of this application.

Signature

*R G C Jenkins & Co*

Date 23 December, 1999

R G C JENKINS & CO

12. Name and daytime telephone number of person to contact in the United Kingdom

Holly Whitlock 0171-931-7141

#### Warning

*After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.*

#### Notes

- If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.
- Write your answers in capital letters using black ink or you may type them.
- If there is not enough space for all relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.
- If you have answered 'Yes' Patents Form 7/77 will need to be filed.
- Once you have filled in the form you must remember to sign and date it.
- For details of the fee and ways to pay, please contact the Patent Office.

Method and apparatus for transmitting a video image

The invention relates to a method and apparatus for transmitting a video image. More particularly, the invention relates to transmission of a video image including an object of interest, such as a face in communications using a mobile video-phone.

In mobile video communication, often the video camera is hand-held and moves relative to the subject. That is particularly the case, for example, in mobile phone based video communication, where the user has to direct a camera linked to the phone handset to point to his own face. This can cause the problem that, because of head and hand movements, the outline of the user's face moves within the frame of the image captured by the camera, possibly even moving outside the frame. One solution for preventing the outline of the head from moving outside the frame is to adjust the focal length of the camera so that the outline of the head occupies a relatively small fraction of the frame. As a result, the probability that the head stays within the frame of the image is increased. However, the resolution of the face image is decreased and so the quality of the video-link is perceived to be poor.

20

US 4,951,140 discloses a device for encoding image data including a face, where the face region is detected and more bits are allocated to the extracted

face region than to the rest of the image. to achieve a better quality image of the face region.

5 According to a first aspect, the invention provides a method of transmitting a video image including an object of interest comprising capturing a sequence of images in which the object of interest occupies a fraction of each image, tracking the object of interest by selecting and extracting a region of each image including the object of interest, and coding only the selected region of each captured image.

10

By arranging that the object of interest occupies a small fraction of the image, the probability that the object of interest stays within the frame of the captured image is increased. At the same time the object of interest occupies a relatively high fraction of the region that is coded. As a result of the invention, the amount of data to be coded is reduced. Even if the object of interest moves around within the frame of the captured image, the extracted region follows the object, so that the object moves less within the extracted region. Preferably, the object of interest is centred in the extracted region, so that the object is stable within the extracted regions.

20

Preferably, the extracted region is displayed at a resolution lower than the resolution of the captured image. Because the object of interest occupies a

relatively large proportion of the terminal display, the user perceives that the quality of the image is improved.

According to a second aspect, the invention provides a method of transmitting  
5 a video image including an object of interest comprising selecting a region of  
an image including the object of interest, the selected region being of a  
predetermined size, and coding the selected region.

As above, the object of interest occupies a relatively high fraction of the  
10 region that is coded.

Preferably, only the selected region is coded and the rest of the captured  
image is discarded. As a result, the amount of information to be coded and  
transmitted is reduced. Preferably, the selected region corresponds to a  
15 predetermined image format having fewer pixels than the capture image of the  
camera. Consequently, the selected region can be coded and displayed using  
known coders and displayed for the known image format without further  
processing to adapt the extracted region to the specific format. This reduces  
the amount of processing required.

20

According to a third aspect, the invention provides a method of transmitting a  
video image including an object of interest comprising selecting a region of

the image greater than the object of interest by a predetermined degree, and coding said region.

5 The third aspect of the invention offers advantages similar to those of the first and the second aspect. Also, more specifically, the size of the selected region varies with the size of the object within the image (which changes for example as the object moves relatively towards or away from the camera) so that the ratio of object data to background data stays approximately the same.

10 According to a fourth aspect, the invention provides a method of operating a video camera comprising arranging the camera so that an object of interest occupies a fraction of the area of the captured image, tracking movement of the object of interest within the captured image, selecting and extracting a region of interest around the object of interest and displaying only the  
15 extracted part of the captured image.

As a result of the invention, it is easier to keep the face of the user within the image captured by the camera while also maintaining a high quality displayed image at the receiving terminal. Also, the amount of information to be coded  
20 may be reduced compared with the prior art. Also, effects of the invention can be obtained using certain standard components such as standard encoders.



Embodiments of the invention will be described with reference to the accompanying drawings of which:

Fig. 1 is a block diagram of a mobile video communication system;

5

Fig. 2 is a block diagram showing the image processing circuit of Fig. 1 in more detail;

Fig. 3 is a block diagram of an image processing circuit in a second  
10 embodiment;

Fig. 4 is a block diagram of an image processing circuit in a third embodiment.

15 An example of an application of the present invention is a mobile video phone communication system. Components of such a system are shown in block diagram form in Fig. 1.

A mobile phone (not shown) includes a camera 2 for capturing images of the  
20 user. The camera 2 is a known type of camera for use in mobile video phones and is part of the phone handset. In an alternative embodiment, the camera is a separate component connected to the phone handset, for example, by a lead

or by wireless communication. The camera digitises images at CIF resolution (352x288 pixels). The optical system of the camera is chosen so that in use the face of the user occupies approximately a predetermined fraction of the target image resolution, which is the resolution of the display 14. Here, the resolution of the display corresponds to QCIF format (176x144 pixels). In this embodiment, the optical system is configured so that in normal use the face occupies approximately 80 % of the target resolution. Of course, the actual fraction of the image occupied by the face of the user will in use depend on various factors, such as the size of the face of the user and where the camera is actually held. Accordingly, the configuration of the camera including the focal length of the optical system is determined on the basis of statistical information representing, amongst other things, the average size of people's faces and what is considered a comfortable distance from the face for holding the camera.

15

The camera is connected to a signal processor 4 for processing signals received from the camera 2 representing the captured image. The signal processor 4 is shown in more detail in Fig. 2. The signal processor includes a face detection module 16, for detecting the size and position of the face or head in the captured image, a face tracking module 18, for tracking the face as it moves in the image, a region selection circuit 20, for selecting a specific region of the image, and a face region extraction module 22. Face-detection

20

circuits and face tracking circuits are known and described, for example, in G. Burel and D. Carel - Detection and Localisation of faces on digital images, Pattern Recognition Letters, 15:963-967, October 1994 and in Lars-Peter Bala, Kay Talmi and Jin Liu - Automatic Detection and Tracking of Faces and Facial Features in Video Sequences, Picture Coding Symposium 1997, 10-12 September 1997, Berlin Germany, the contents of which are incorporated by reference. The signal processor 4 operates to select and extract a desired region of the image including the face region, as will be described in more detail below.

10

An output of the signal processor 4 is connected to an encoder 6, for encoding signals representing the extracted region of the image signal. The encoder 6 is a known encoder. The encoder is connected to a transmitter 8, for transmitting the coded signal in a known manner:

15

The receiving side of the system is a receiving terminal in the form of a second mobile phone (not shown). The second phone includes a known receiver 10 for receiving the transmitted signal, a decoder 12 connected to the receiver for decoding the received signal, and a display 14 for displaying the received image in QCIF format.

20

In operation, an image is captured by the camera 2, and the resulting signals are input to the signal processor 4. The image is analysed by the face-detection module 16, which determines the position and size of the face within the image in a known manner.

5

Information regarding the location and size of the face are input from the face-detection module 16 to the region selection circuit 20, which determines the size and location of the window to be selected from the main image. In this embodiment, the region selection circuit 20 is configured to select a window of a predetermined size centred on the face. More specifically, the region selection circuit selects a window having the same resolution as the display. Thus, in this case, the region selection circuit is configured to select a region sized 176x144 pixels, centred on the face region. The centre can be defined and determined in any suitable manner. In this embodiment, the centre of the face is the mid-point based on the extremes vertically and horizontally of the flesh-region.

15

Because of the set up of the optical system of the camera 2, as explained above, the face region occupies approximately 80% of the selected window (in normal use), that is approximately 100x150 pixels. Thus, in normal operation, assuming the face is in the centre of the CIF image, there is a boundary around the selected region of 126 pixels in the vertical direction and

20

69 pixels in the horizontal direction. Thus, even if the outline of the face is displaced horizontally or vertically because of head and/or camera movements, it will still be reflected within the CIF image as long as displacement is less than the distances mentioned above. For the above  
5 example, the vertical face displacement in the image plane can be 1.26 times the face width. To achieve a similar coverage in a conventional system with QCIF resolution, the width of a face would have to be 50 pixels.

The face region extraction module 22 receives signals from the camera and  
10 from the region selection circuit and extracts the window including the face region from the image from the camera. The extracted window is then transferred to the standard QCIF coder 6 for coding using a suitable known coding method. The remainder of the image is discarded. These steps are performed for each frame of the captured video images, the face being tracked  
15 by the tracking circuit to reduce the amount of processing required.

As the face-region moves within the captured image, the extracted window also moves around the captured image. Because the face detection module is supported by the face tracking module, it is not necessary to do a full face-  
20 detection process in each frame and thus the amount of processing is reduced. Because the region selection circuit 20 is configured to select the window centred on the face region, the face is stabilised within the extracted window.

Thus, even if the head moves within the captured image, it does not move within the extracted window. This is less distracting for the viewer.

5 The coded signal is transmitted, received and decoded by the receiving terminal, which displays the image of the face in QCIF format. Because of the process of selection of a region of the captured image which has the face in the middle of the region and which is of QCIF resolution, the displayed image has the face in the middle and is the correct resolution for the display. Also, the face is displayed as a higher fraction of the image than in the  
10 captured image, which gives the impression of better resolution.

A second embodiment of the invention will be described with reference to Fig. 3. The second embodiment corresponds to the first embodiment but has a  
15 region extraction and scale module 24 in place of the region extraction module 22.

In embodiment 2, a region surrounding a face-region is selected as in the first embodiment. However, the extracted region is also scaled to compensate for  
20 variations in the size of the face region resulting from movements of the user relatively towards and/or away from the camera. In other words, the extraction and scale module 24 also performs a digital zoom procedure on the extracted

region. Scaling is performed so that the fraction of the extracted region occupied by the face region is approximately the same in each successive extracted region. The size of the extracted region after scaling in pixels is the same as before scaling. Coding, transmission and display are carried out as in  
5 the first embodiment. The second embodiment has the advantage that it results in a more stable image that is less distracting to the viewer.

A third embodiment of the invention will be described with reference to Fig.  
4.

10

The third embodiment corresponds to the first embodiment, subject to modifications to the region selection circuit and the region extraction module. Also, the region selection circuit 20' has a user input.

15 In this embodiment, the region selection circuit 20' operates to select a region around the face region such that the face region occupies a predetermined fraction of the selected region. The predetermined fraction is selected by the user by way of the user input, in the form of a keyboard. In an alternative  
20 embodiment, the fraction may be fixed by the manufacturer. In this example, it has been selected that the face region occupies 80% of the selected region. In other words, the size of the selected region is 125% of the face region.

The face detection and tracking modules 16, 18 detect and track the face region as in embodiments 1 and 2. The region selection circuit 20' then selects a region around the face region in accordance with the preferences. Here, the region is a rectangular region, scaled in relation to the face region, and centred on the face.

The selected region is then extracted by the region extraction module 22'. The size of the extracted region in pixels is dependent on the size of the face in the captured image, and it may vary, for example, as the head moves closer to or further away from the camera. Thus, the selected region is scaled to a predetermined size in the region extraction module. Here, the region is scaled to QCIF format, so that it can then be coded using a standard QCIF encoder 6. Alternatively, the captured image can be subjected to digital zoom before the face region is extracted so that the size of the face, and hence the size of the extracted region, is the same in each frame.

Subsequently, the coded signal is transmitted and displayed as described above.

The above embodiments have been described in relation to mobile video phone communication. The invention may also be used in other applications, such as in video-conferencing and transmission of video images from cameras



connected to personal computers. The embodiments describe selection of a region including the face of a speaker as an object of interest, but the invention can be applied in relation to any other object of interest. The invention has been described using CIF and QCIF, but other formats may be used. In embodiment 3, instead of selecting a region that is a certain percentage greater than the face-region, the selected region could be a predetermined amount greater, for example, longer and wider by a certain number of pixels.

Claims

1. A method of transmitting a video image including an object of interest comprising capturing a sequence of images in which the object of interest occupies a fraction of each image, tracking the object of interest by selecting and extracting a region of each image including the object of interest, and coding only the selected region of each captured image.  
5
2. A method as claimed in claim 1 comprising stabilising the object of interest within the extracted region.  
10
3. A method as claimed in claim 2 wherein the extracted region is selected so that the object of interest is centred within the extracted region.
- 15 4. A method as claimed in any one of claims 1 to 3 comprising transmitting the coded region, and decoding and displaying the selected region.
5. A method as claimed in claim 4 wherein the extracted region is displayed in a format comprising fewer pixels than the format of the captured image.  
20

6. A method as claimed in any one of claims 1 to 5 in which the object of interest occupies less than a predetermined fraction of each image.

7. A method as claimed in any one of claims 1 to 5 in which the object of interest occupies a small fraction of each image.

8. A method of processing a video image including an object of interest comprising selecting a region of an image including the object of interest, the selected region being of a predetermined size, and coding the selected region.

10

9. A method as claimed in claim 8 wherein only the selected region is coded and the rest of the captured image is discarded.

10. A method as claimed in claim 8 or claim 9 wherein the selected region corresponds to a predetermined image format having fewer pixels than the format of the image capture of the camera.

15

11. A method as claimed in claim 10 wherein the captured image is in CIF format and the selected region is in QCIF format.

20

12. A method as claimed in any one of claims 8 to 11 wherein the selected region is scaled to compensate for movements of the object of interest backwards and forwards relative to the camera.

5 13. A method as claimed in any of claims 8 to 12 wherein the object of interest is stabilised within the selected region.

14. A method as claimed in claim 13 wherein the selected region is such so that the object of interest is centred in the selected region.

10

15. A method of processing a video image including an object of interest comprising selecting a region of the image including the object of interest and which is greater than the area occupied by the object of interest by a predetermined degree, and coding said region.

15

16. A method as claimed in claim 15 wherein the object of interest occupies a predetermined percentage of the selected region.

17. A method as claimed in claim 15 or claim 16 comprising scaling the  
20 selected region to a predetermined size.

18. A method as claimed in claim 17 wherein the predetermined size corresponds to a known format.

19. A method as claimed in claim 18 wherein the captured image is in CIF  
5 format and the extracted region is scaled to QCIF format.

20. A method of transmitting video images comprising processing video  
images according to a method as claimed in any one of claims 1 to 19,  
transmitting the encoded image data, and receiving, decoding and displaying  
10 the image data.

21. A method of operating a video camera comprising arranging the  
camera so that an object of interest occupies a fraction of the area of the  
captured image, tracking movement of the object of interest within the  
15 captured image, selecting and extracting a region of interest around the object  
of interest and displaying only the extracted part of the captured image.

22. An image processing circuit comprising means for extracting a region  
of each image including an object of interest and coding only the selected  
20 region of each captured image.

23. An image processing circuit comprising means for selecting a region of an image including an object of interest, the selected region being of a predetermined size, and coding the selected region.

5 24. An image processing circuit comprising means for selecting a region of the image such that the object of interest occupies a predetermined percentage of the region, and for coding said region.

25. A video image processing circuit comprising means for performing a  
10 method as claimed in any one of claims 1 to 21.

26. A video image processing device comprising a camera and a circuit as claimed in any one of claims 21 to 25.

15 27. A mobile phone comprising a circuit as claimed in any one of claims 21 to 25 or a device as claimed in claim 26.

28. Apparatus for processing video images substantially as hereinbefore described as an embodiment with reference to the respective accompanying  
20 drawings.

29. A method of processing video images substantially as hereinbefore described as an embodiment with reference to the respective accompanying drawings.

Method and apparatus for transmitting a video imageAbstract

5 A method of transmitting a video image including an object of interest comprises capturing a sequence of images in which the object of interest occupies a fraction of each image, tracking the object of interest by selecting and extracting a region of each image including the object of interest, and coding only the selected region of each captured image.

10

Fig. 1



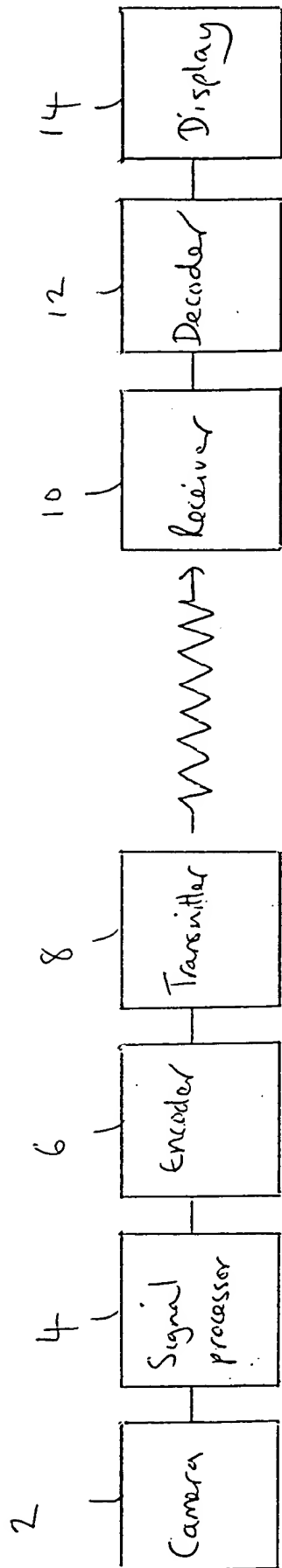


Fig. 1.

**THIS PAGE BLANK (USPTO)**

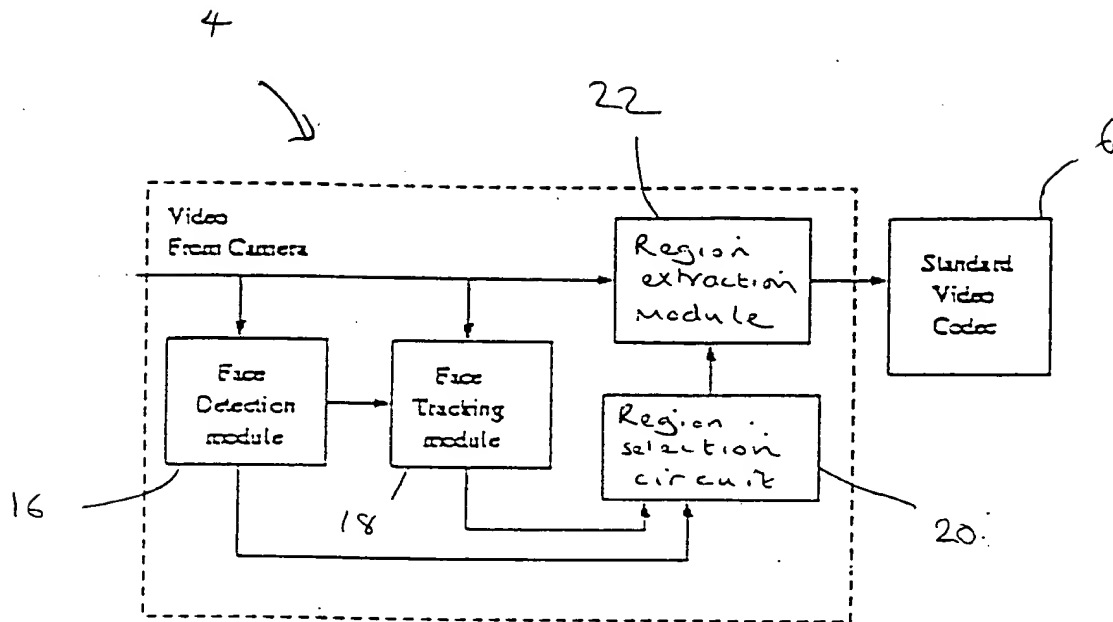


Fig. 2

THIS PAGE BLANK (USPTO)

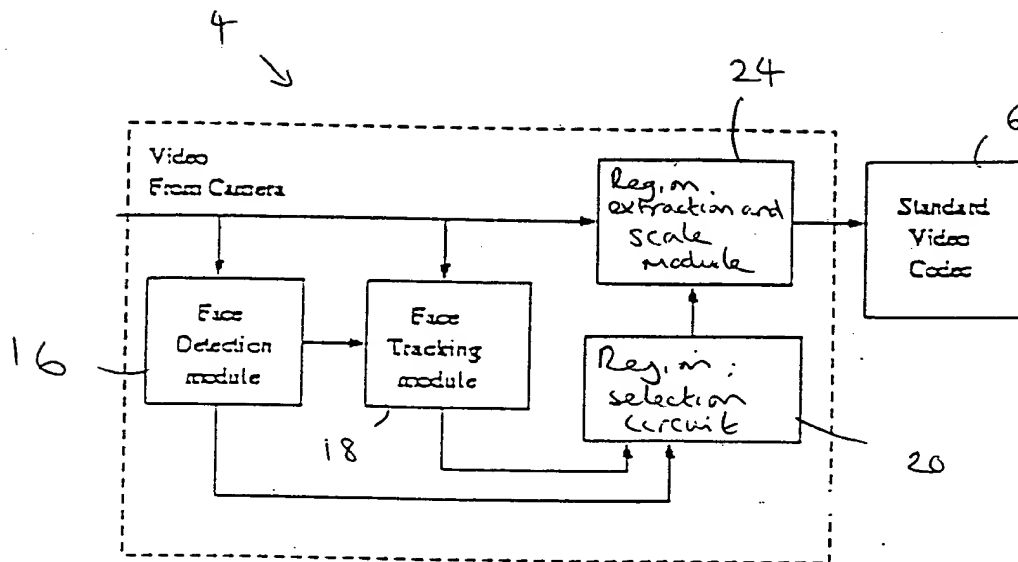


Fig 3

**THIS PAGE BLANK (USPTO)**

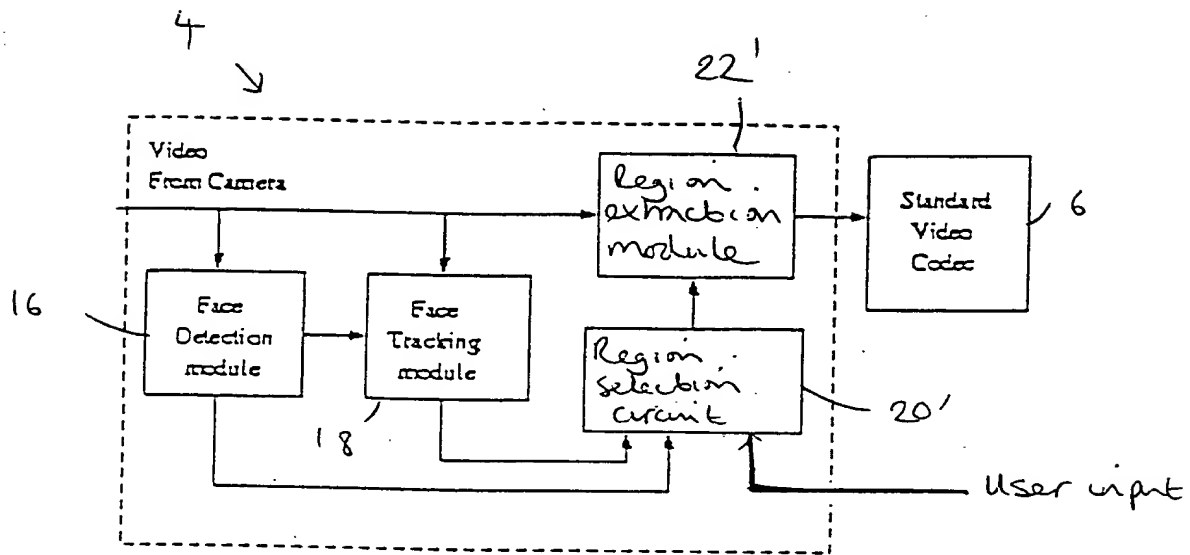


Fig 4

THIS PAGE BLANK (USPTO)



**22850**

SERIAL NO.: 09/785,435

FILING DATE: FEB. 20, 2001